

## **Evaluating End-User Services: Success or Satisfaction?**

**by Melvon L. Ankeny**

Does an indication of "satisfaction" by surveyed end users of online information services signify that they've actually been successful in obtaining needed information? This article reviews past evaluations of online search services and, from a study of 600 untrained end users, draws comparisons between users' satisfaction and their relative success.

Online databases are now recognized as standard reference tools for the librarian in answering information requests. Today the direct use of online services by end users is also becoming standard practice. As more end-user services become available in libraries, studies evaluating their use are being published in the literature. A major focus of past studies has been the measurement of users' satisfaction with mediated and end-user online searches. Some recent studies have raised concerns that, satisfaction levels aside, the actual success of end users in finding relevant information may be quite low.

This article compares two methods of assessing the success of end-user searches. The first is the commonly used method of rating the satisfaction with or utility of a search through a response to a single question that uses a Likert-type scale, generally 1 to 5 points (see Figure 1). The other is a more discriminating method that uses responses to 3 questions using a 6-point anchored scale for success in conjunction with a 3-point scale for satisfaction, and a section listing 10 possible reasons for dissatisfaction (see Figure 2).

### **Literature Review**

Earlier studies on user satisfaction with mediated online search services (as reported, for example, by Kobelski and Trumbore)<sup>1</sup> were called into question in a 1983 article by Auster.<sup>2</sup> Auster stated that the overwhelming percentage of users reporting satisfaction with their searches "raises questions not only about the level of expectations and sophistication of current users but about the discriminating power of instruments used to measure their satisfaction." Auster and Lawton, in their own study of user satisfaction, used a more discriminating instrument where the satisfaction variable was measured by responses to five items.<sup>3</sup>

The evaluation of end-user online services has mirrored in part that of mediated online services. One of the earlier studies of end users by Sewell and Teitelbaum reported on successes and problems of nearly 4,000 searches on MEDLINE and other National Library of Medicine (NLM) databases.<sup>4</sup> This 1982 study looked at levels of search complexity and sophistication, patterns of searching and printing, and degrees of database usage for end-user searches performed between 1976 and 1980. Preliminary observations included findings of "missed opportunities" in 55 percent of MEDLINE searches using drug or chemical terms. A 1983 paper by Hurt looked at differences in levels of user satisfaction for mediated searches and end-user searches, reporting a significantly higher level of satisfaction for mediated searches.<sup>5</sup> Satisfaction

was measured by the end user on a six-point scale. Further testing indicated that "the major reason for low satisfaction was difficulty with protocol."

Janke reported on a 1983 survey of 25 end users of BRS/After Dark that focused on the issue of relevancy.<sup>6</sup> In all, 57.1 percent of the respondents claimed that at least 60 percent of the references they retrieved were relevant to their research. In a follow-up survey reported in 1984, Janke stated that "perhaps the most fascinating generalization is this: in both surveys approximately half of the respondents stated that 60 percent or more of the references they retrieved were highly relevant to their research needs, while the other half claimed that 30 percent or less achieved this degree of relevance."<sup>7</sup> This led him to conclude that "at least 50 percent of end users have difficulty getting satisfactory results." Janke also found that the end-user system in question had corrected the protocol obstacle cited by Hurt and reported that 65 percent of the users found the system "extremely easy to use" or "easy to use."

Trzebiatowski also reported on a 1983 study of first-time end users of BRS/After Dark.<sup>8</sup> Twenty end users whose reference questions "seemed appropriate" for database searching were chosen for participation. Each end user was given a pre-search orientation of approximately 30 minutes. When asked to characterize their search results, 40 percent called them excellent, 45 percent good, and 15 percent satisfactory. When asked about the relevancy of their retrieved citations, 20 percent claimed that 75-100 percent were useful; 35 percent claimed 50-75 percent were useful; 20 percent claimed 25-50 percent were useful; and 25 percent claimed fewer than 25 percent were useful. End users reported encountering problems with the terminal 30 percent of the time, the commands 35 percent of the time, and the search strategy 50 percent of the time. Trzebiatowski concluded that the infrequent end user in an academic setting would continue to need instruction and that the role of the intermediary as consultant would be part of the "successful search."

Slingsluff, Lev, and Eisan reported on the use of BRS/After Dark in an academic health sciences library, observing and recording high levels of user satisfaction by means of informal exit interviews.<sup>9</sup> Problems identified were "almost exclusively related to system mechanics" and "almost all searchers indicated that they were satisfied with their results." The end user's frame of reference for "satisfaction" in this instance was a comparison with results which would have been obtained manually rather than with results of a search by a trained intermediary. Their "satisfaction" with searches was also demonstrated by their repeated usage of the service.

**Figure 1**  
**First Study: Survey Questionnaire (Selected Questions)**

---

1. Did you obtain the information you wanted?				
1. Yes				
2. No				
2. Overall, how would you rate these online services?				
1	2	3	4	5
Low				High

---

Sewell and Teitelbaum's 1986 article reported on a number of studies they had conducted during 11-plus years of work with end users of NLM databases.<sup>10</sup> Among the questions they set out to answer was: What are the major factors in end users' success or failure in finding the information they need? Their work included a base-line study of all search session transaction logs (1976-1984), a retrospective study of 37 repeat searchers, a corroborating survey questionnaire in 1981, and a follow-up study (July-December 1982) in which they reviewed



to the amount of connect time and the calculated cost-effectiveness of the search. Walker concluded that information levels would only be minimally affected by moving to end-user searching but that the cost implications of longer search times would require consideration by information managers.

## Overview

Online end-user services began at the Ohio State University Business Library in the summer of 1987 with free access to DIALOG Business Connection and Dow Jones News/Retrieval. Business Connection is a menu-driven system accessing a number of company- and product-focused databases from the full DIALOG service. OSU's Dow Jones News/Retrieval package offers menu-driven access to the Dow Jones business information databases with the exception of the Text services and selected high-cost databases.

Because of staffing configurations at the Business Library, the online services were initiated with very little instruction for the end users. When they signed on at the circulation desk, end users were provided with a two-page set of instructions covering both services.

**Table 1 First Study: Responses to: Did you obtain the information you wanted?**

Response	Number	% of Total*
Yes	148	77.9%
No	42	22.1%
*Due to rounding, percentages in this and following tables may not add to 100.		

**Table 2**

**First Study: Responses to: Overall, how would you rate these online services?**

Rating	Number	% of Total*
1 (Low)	2	1.1%
2	5	2.6%
3	28	14.7%
4	70	36.8%
5 (High)	85	44.7%

## First Study

To gather information about usage patterns and to determine whether the services were of value to the end users, a simple one-page questionnaire was given to all end users over a four-month period in late 1987. End users were queried as to the system chosen, their preference for one or the other, whether they had found the information they were seeking, and what overall rating they would assign to OSU's online end-user services. Of 190 end users surveyed in this first study, 77.9 percent answered "yes" and 22.1 percent answered "no" to "Did you obtain the information you wanted?" (see Table 1). On the overall rating of the services, 44.7 percent of the end users rated them as 5 and 36.8 percent rated them as 4 on a scale of 1 to 5 (see Table 2). The 77.9 percent of all end users indicating that they were successful in finding the information they wanted compares favorably to the 81.5 percent who gave the services an overall rating of 4 or 5 on the five-point scale.

Findings from the first study indicated that untrained end users were reasonably successful in finding what they wanted. While these results were gratifying, they did not conform to the findings of Kirby and Miller, which suggest that untrained end users should have considerably lower success rates.<sup>14</sup> Consequently, there was concern that the reports of success and satisfaction obtained from the first study's questionnaire might have been inflated. If this were so, the first study might fail to identify or suggest solutions to problems in end-user operations.

## Second Study

A second and larger study was developed using the same measure of success as was used for traditional reference service in the Wisconsin-Ohio Reference Evaluation Program (see Figure 2).<sup>15</sup> This patron form has been tested and found to have a reliability coefficient of .90. In addition, this success measure has been found to have external validity—i.e., findings obtained with this questionnaire are not significantly different from those gained through unobtrusive observation. This form has now been used in conjunction with evaluative studies involving over 7,000 patrons in 100 libraries. Control against positive bias or inflation has been exercised by a rigorous scoring system. For a search to qualify as a complete success, three conditions must be met: the patron must indicate that s/he found *exactly* what was wanted, that s/he was *fully* satisfied with the search, and that s/he marked none of the 10 listed reasons for dissatisfaction. This format allows the patron three opportunities to reflect on the outcome of the transaction.

Six hundred responses to this second questionnaire were collected over a three-month period at the end of 1987 and beginning of 1988. The questionnaires were distributed when patrons signed on to use the end-user systems and collected at the end of each session. Not all end users complied with the request to complete a questionnaire and no attempt was made to differentiate between patrons who had previous experience with either of the end-user systems and those who had not.

## Results and Discussion

Of the 600 searches in the sample, 233 met all three criteria for complete success and 367 were less than successful, yielding an overall success rate of 38.8 percent (see Table 3). This rate closely resembles the 40 percent success rate reported by Kirby and Miller, who used a different methodology and a smaller sample of untrained end users on BRS/Saunders Colleague.<sup>16</sup>

In comparing the first and second studies at OSU, it can be seen that the comparable questions "Did you obtain the information you wanted?" (first study), and "Did you find what you searched for?" (second study) yielded much different results. Specifically, the second form of the question, with its six options for response, yielded a much lower report of success (42.8 percent) than the first study's "yes-no" question (77.9 percent). Requiring that the definition of success also include full satisfaction allows for the cases of patrons who find exactly what is wanted but feel dissatisfied with some component of the search, e.g., they experienced difficulties with the system, they retrieved too much or too little information (see Table 4). When this criterion for success is added, success drops to 38.8 percent.

A comparison of responses received on the second study's five-point Likert scale rating for online services (question 4, Figure 2) to responses received on the same study's multiple-choice questions measuring relative success (questions 1-3, Figure 2) suggests that different

concepts were being measured. For example, Table 5 reveals that 447 or 75.8 percent of the 590 end users responding to the question on ratings gave an overall rating of 4 or 5 to OSU's online services. However, only 209 of these 447 searches were successful by the three-criteria success measure. If one considers only the 38.8 percent of end users who gave online services the *highest* rating, however, this proportion corresponds exactly to the 38.8 percent of the complete 600 sample who were deemed to have fully successful searches.

### **Reported Reasons for Dissatisfaction**

Of the 367 less-than-successful searches, 257 respondents checked at least one item under question 3 (If not completely satisfied with the information found, why not?) (see Table 4). Of these 257 respondents, 18.7 percent reported that they experienced problems using the system or found it too difficult to use (compared to Kirby and Miller's study, where 22.6 percent of failures were attributed to system problems). Users identified problems with connecting to the system via the communication software and printing results; they also expressed the need for better printed instructions and more informed and accommodating staff. Although it was not possible to track individual searches within a system to determine database selection error, 3.9 percent of the remaining searches listed search subjects determined by the author to be inappropriate topics for the system chosen (i.e., information sought was not available on the system). The other 77.4 percent of respondents indicated dissatisfaction with the amount, relevancy, or level of the information retrieved,

In the absence of transactions logs (with complete statements of the subject terms used and the sequence of search procedures followed), it is not possible to reconstruct the end-user searches; thus only limited conclusions as to reasons for failure can emerge. Neither system requires the use of Boolean logic (the Text portions of Dow Jones News/Retrieval were not available for access), so some of the classic difficulties with search strategy were not of concern. However, each system has its own particular difficulties and the reliance on a menu-driven system sometimes results in an unwieldy process ultimately leading nowhere.

### **Problems with Online Searches**

Possible reasons for failure, drawn from repeating less-than-successful searches using the indicated search topics, correspond in part to problems that are experienced using print sources. This study revealed that a major recurring problem with the Business Connection is the confusing information provided in response to requests for corporate intelligence. Because the service draws from a number of databases in which company name and location information may vary, the end user receives a list of variant names that must be tracked through each database's menu to determine if any of the listings contain useful information. The addition of an online stock symbol directory to the company finder database of Business Connection would be helpful (as would some way for the system to differentiate between stock symbols—e.g., LTD for The Limited—and portions of company names—e.g., Labatt Ltd.), as this avenue of access in most cases leads directly to the main listing of information on publicly traded companies.

**Table 3 Second Study: Responses to Success Measures**

Response	Number	% of Total
Found exactly (Question 1)	257	42.8%
Fully satisfied (Question 2)	352	58.7%
No reason for dissatisfaction (Question 3)	343	57.2%
Responses fulfilling all three success criteria	233	38.8%

**Table 4****Second Study: Responses to: If not completely satisfied with the information found, why not?**

Source of Dissatisfaction	No. of Respondents	% of Total
System problems/ Too difficult to use	48	18.7%
Inappropriate Topic: Information not available on system searched*	10	3.9%
Other Reasons for dissatisfaction	199	77.4%

\* Author judged 10 searches to be inappropriate for the system chosen (i.e., information sought was not available on system).

**Table 5****Second Study: Responses to: Overall, how would you rate these online services?\***

Rating	Total		Less than Successful		Successful	
	No.	%	No.	%	No.	%
1 (Low)	27	4.6%	12	2.0%	15	2.5%
2	26	4.4%	23	3.9%	3	.5%
3	90	15.3%	86	14.6%	4	.7%
4	218	36.9%	164	27.8%	54	9.2%
5 (High)	229	38.8%	74	12.5%	155	26.3%

\* All percentages are determined on the basis of 590 respondents in the sample who rated the services.

Problems associated with form and spelling of names, geographic location of parent company, and possibility of stock symbol access suggest that end users would be well advised to set a standard reference tool to determine as much information as they can before going to the online source. A major point of access to product and industry information on Business Connection is the use of the Standard Industrial Classification (SIC) number and the related Predicast code. Choices must be made whether to enter an SIC or Predicast number, or to search with a generic product name. Choice of access determines the varying results. End users are often unsure of the structure of the classification schemes and in many cases use only a very specific number when a broader, more encompassing number would also yield relevant information. Usage of these numbers is complicated by the fact that some databases continue to use the previous edition of the SIC and the end user must search using two or more numbers if they were changed from edition to edition. End users seeking a market share report may also be confused by menu prompts to enter either an SIC or Predicast code, when only the four-digit SIC number will generate a market share report from the Trinet Establishment Database.

## Solutions

To help solve some of these searching problems, librarians may want to provide end users with a copy of the latest edition of the SIC Manual (with reference to the tables in the back

of this edition showing changes from the previous edition), a copy of the Predicast classification scheme, and alphabetic product index from the Predicast F&S Index.<sup>17</sup> In addition, they may want to develop easily comprehensible search aids (focusing on trouble-shooting aids) that target problem areas such as system mechanics and search instructions, and provide recommendations for determining company intelligence offline prior to searching (to determine name form, stock symbol, and geographic location). Addressing these specific concerns will undoubtedly lead to higher levels of success for the end user.

## Conclusions

Evidence is accumulating that actual success rates of end-user searches are quite low, which substantiates Janke's original observation in 1984 that at least 50 percent of end users experience difficulty.<sup>18</sup> The success rate of the end users in this study (38.8 percent) is similar to that found in Kirby and Miller's study of BRS/Saunders Colleague (40 percent).<sup>19</sup> This study appears to indicate that in many cases high levels of reported end-user "satisfaction" with computerized services may not reflect true success rates. One factor that appeared to result in inflated success ratings was the use of a two-point (yes-no) success scale rather than a multiple-choice success scale, such as that used in the Wisconsin-Ohio Reference Evaluation Program. In reference to the rating scale, it also appears that on a five-point Likert scale (without brief phrases to anchor each scale point), a large amount of inflation occurs in the area of the fourth rating, while the fifth (or top) rating appears relatively free of inflation.

These findings are in line with concerns raised by Auster about the ability of some types of evaluation instruments to measure the true success of those engaged in computerized searches.<sup>20</sup> Although more research needs to be done, the results of this study suggest that users can make critical judgments about the results of their own searches if given an evaluation instrument with sufficient distinctions for satisfaction and success. Needed now are further studies in other settings to determine cross-service, cross-institution rates of success/failure—as a means of both targeting areas for improvement and placing particular services in perspective.

## References

- <sup>1</sup> Pamela Kobelski and Jean Trumbore, "Student Use of Online Bibliographic Services," *Journal of Academic Librarianship* 4 (March 1978): 14-18.
- <sup>2</sup> Ethel Auster, "User Satisfaction with the Online Negotiation Interview: Contemporary Concern in Traditional Perspective," *RQ* 23 (Fall 1983): 47-59.
- <sup>3</sup> Ethel Auster and Stephen B. Lawton, "Search Interview Techniques and Information Gain as Antecedents of User Satisfaction with Online Bibliographic Retrieval," *Journal of the American Society for Information Science* 35 (March 1984): 90-103.
- <sup>4</sup> Winifred Sewell and Sandra Teitelbaum, "Preliminary Observations of Nonmediated Search Behavior of Pathologists and Pharmacists," *American Society for Information Science—Proceedings of the Annual Meeting* 19 (October 1982): 276-278.
- <sup>5</sup> C.D. Hurt, "Intermediaries, Self-searching and Satisfaction," *National Online Meeting Proceedings—1983* (Medford, NJ: Learned Information, 1983): 231-237.
- <sup>6</sup> Richard V. Janke, "BRS/After Dark: The Birth of Online Self-Service," *Online* 7 (September 1983): 12-29.
- <sup>7</sup> Richard V. Janke, "Online After Six: End User Searching Comes of Age," *Online* 8 (November 1984): 15-29.
- <sup>8</sup> Elaine Trzebiatowski, "End User Study on BRS/After Dark," *RQ* 23 (Summer 1984): 446-150.
- <sup>9</sup> Deborah Slingluff, Yvonne Lev, and Andrew Eisan, "An End User Search Service in an Academic Health Sciences Library," *Medical Reference Services Quarterly* 4 (Spring 1985): 11-21.
- <sup>10</sup> Winifred Sewell and Sandra Teitelbaum, "Observations of End-User Online Searching Behavior Over Eleven Years," *Journal of the American Society for Information Science* 37 (July 1986): 234-245.



<sup>11</sup>"Martha Kirby and Naomi Miller, "Medline Searching on Colleague: Reasons for Failure or Success of Untrained End Users," *Medical Reference Services Quarterly* 5 (Fall 1986): 17-34; and Martha Kirby and Naomi Miller, "Medline Searching on BRS Colleague: Search Success of Untrained End Users in a Medical School and Hospital," *National Online Proceedings—1985* (Medford, NJ: Learned Information, 1985): 255-263.

<sup>12</sup>Alice C. Littlejohn, "End-User Searching in an Academic Library—The Students' View," *RQ* 26 (Summer 1987): 460-66.

<sup>13</sup>Geraldene Walker, "The Search Performance of End-Users," *National Online Meeting Proceedings—1988* (Medford, NJ: Learned Information, 1988): 403-410.

<sup>14</sup>Kirby and Miller, "Medline Searching on Colleague" and "Medline Searching on BRS Colleague."

<sup>15</sup>Marjorie E. Murfin and Gary M. Gugelchuk, "Development and Testing of a Reference Transaction Assessment Instrument," *College & Research Libraries* 48 (July 1987): 314-338.

<sup>16</sup>"Kirby and Miller, "Medline Searching on Colleague" and "Medline Searching on BRS Colleague."

<sup>17</sup>U.S. Office of Management and Budget, *Standard Industrial Classification Manual* (Springfield, VA: National Technical Information Service, 1987); and *Predicasts F&S Index United States* (Cleveland, OH, Predicasts, 1979-).

<sup>18</sup>Janke, "Online After Six," p. 20.

<sup>19</sup>"Kirby and Miller, "Medline Searching on Colleague" and "Medline Searching on BRS Colleague."

<sup>20</sup>Auster, "User Satisfaction."